In-silico analysis and implementation of a multicellular feedback control strategy in a synthetic bacterial consortium

Gianfranco Fiore,[†,‡,⊥] Antoni Matyjaszkiewicz,[†,‡,⊥] Fabio Annunziata,[¶,‡] Claire Grierson,[§,‡] Nigel J. Savery,[¶,‡] Lucia Marucci,[*,†,‡,#] and Mario di Bernardo[*,†,‖,‡,#]

Department of Engineering Mathematics, University of Bristol, Merchant Venturers' Building, Woodland Road, Bristol BS8 1UB, UK, BrisSynBio, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK, School of Biochemistry, University of Bristol, Biomedical Sciences Building, University Walk, Bristol BS8 1TD, UK, School of Biological Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK, and Department of Electrical Engineering and Information Technology, University of Naples Federico II, 80125, Naples, Italy

E-mail: lucia.marucci@bristol.ac.uk; m.dibernardo@bristol.ac.uk

Abstract

Living organisms employ endogenous negative feedback loops to maintain homeostasis despite environmental fluctuations. A pressing open challenge in Synthetic Biology

*To whom correspondence should be addressed
†Department of Engineering Mathematics, University of Bristol
‡BrisSynBio, University of Bristol
¶School of Biochemistry, University of Bristol
§School of Biological Sciences, University of Bristol
‖Department of Electrical Engineering and Information Technology, University of Naples
⊥Contributed equally to this work as first authors
#Contributed equally to this work

is to design and implement synthetic circuits to control host cells' behavior, in order to regulate and maintain desired conditions. To cope with the high degree of circuit complexity required to accomplish this task and the intrinsic modularity of classical control schemes, we suggest the implementation of synthetic endogenous feedback loops across more than one cell population. The distribution of the sensing, computation and actuation functions required to achieve regulation across different cell populations within a consortium allows the genetic engineering in a particular cell to be reduced, increases the robustness, and makes it possible to reuse the synthesized modules for different control applications. Here, we analyze, *in-silico*, the design of a synthetic feedback controller implemented across two cell populations in a consortium. We study the effects of distributing the various functions required to build a control system across two populations, prove the robustness and modularity of the strategy described and provide a computational proof-of-concept of its feasibility.

Keywords

synthetic biology, *E. coli*, synthetic microbial consortia, feedback control, gene networks, mathematical modeling

The majority of living organisms can maintain homeostasis despite external stimuli or environmental fluctuations. Hormone secretion and signaling pathways functioning in multicellular organisms (*1*), as well as the control of bacterial chemotaxis (*2*), are only a few of the many examples in which the internal state of biological systems is regulated or maintained by employing a negative feedback mechanism.

In the last few years, experimental approaches to implement exogenous negative feedback control schemes have been proposed in order to achieve real-time control of gene expression

in living cells (*3–8*). In all these applications, the experimental setup comprises a device to grow cells, a sensing apparatus (cytofluorimeter or fluorescence microscope) to monitor cells' behavior (quantify fluorescence from fluorescent reporters), a PC running control algorithms and a set of actuators to provide inputs to the cells according to the control objective and their current status.

An open challenge in Synthetic Biology is to synthesize endogenous negative feedback controllers, where all the sensing, computation and actuation functions are embedded in living cells. The range of possible applications of synthetic regulators can span from the optimization of chemical production in bioreactors (*9*, *10*), to targeted drug delivery in multicellular organisms (*11*).

To address this challenge, the implementation of synthetic feedback control schemes in single cells has been proposed. Typical control goals are set-point regulation (*12*, *13*), or signal tracking control (*14*) of target proteins. These complex functions are achieved through the design of generally complicated synthetic circuits that can be difficult to characterize and integrate in a single cell, often resulting in metabolic burden that can be self defeating for the host (*15*, *16*). Moreover, once cells have been synthetically engineered, any change in the control strategy or its application requires re-engineering the entire system leading to poor modularity and adaptability of the original design and its parts (*17*).

In order to overcome these drawbacks, inspired by the intrinsic modularity of classical control schemes, we propose to distribute the sensing, computation and actuation functions across different cell populations within a microbial consortium. Recently, the construction and the study of a synthetic oscillator implemented across two distinct cell colonies has clearly shown how the genetic engineering of interacting microbial populations can be exploited to achieve complex and robust population-level behaviors (*18*). Indeed the interaction of microbial populations can be advantageous in accomplishing complicated tasks better than a single population can do, whilst beneficially guaranteeing increased robustness to environmental fluctuations (*19*, *20*). Moreover, the engineering of synthetic microbial consortia

makes it possible to physically separate the components and modules required to achieve the desired functions, hence reducing the unwanted effects of retroactivity in biological circuits (as defined in (*21*)) by which standard parts available in Synthetic Biology can significantly change their behavior upon interconnection (*22*, *23*).

The aim of this paper is to present a thorough *in-silico* proof-of-concept analysis and design of a multicellular feedback controller in a consortium of two interacting populations where one (the 'Controllers') has the task of monitoring and controlling in real time the concentration of a target molecular species in the other (the 'Targets'). We analyze the proposed implementation via two different computational approaches, one consisting of ODE-based models and the other implemented as an agent-based simulation of the consortium. After presenting the results of the *in-silico* model, we provide hints for the *in-vivo* realization of the proposed multicellular control scheme.

Results and Discussion

Multicellular feedback control design

The goal of a control strategy is to make the output of a system of interest follow a desired reference signal (*24*). An example is that of temperature regulation, where feedback control is employed to maintain temperature at some desired constant value, or autopilots that need to make planes track a desired route. The structure of a classical feedback control architecture consists of different functional blocks interconnected with each other, each performing a specific function required to achieve the control goal. Specifically, a feedback control scheme comprises: *a)* a sensing apparatus for measuring the output of the system whose behavior is regulated, *b)* a comparator needed to quantify the *control error*, in terms of the mismatch between the measured output and the desired reference value, say $r(t)$, *c)* a computational device (the controller or regulator) able to take a control decision and determine the action to be taken in order to minimize the control error and *d)* a set of actuators, receiving the

control decision calculated by the controller, and turning it into a physical action on the system being controlled in order to steer its dynamics towards the reference.

In our proposed implementation these different functions are distributed across two interacting microbial populations, where specifically one population (the 'Controllers') contains synthetic circuits to sense and control the status of a process in the other (the 'Targets'). The Controllers can receive an external signal (e.g., an inducer molecule) so that the desired reference level of the process to be regulated in the Target cells can be set (Figure 1a). The two populations communicate with each other through the control input $\hat{u}$ (coming from the Controllers to the Targets), and the process output $\hat{y}$ (fed back from the Targets to the Controllers).

A schematic biological implementation of the feedback control strategy is depicted in Figure 1b. Here, communication within the synthetic consortium is achieved via the release of signaling molecules into the growth medium. In the Controllers, an external reference signal inhibits the production of the species A that, in complex with the signaling molecule Q_2, generates B. The latter, in turn, catalyzes the synthesis of another signaling molecule, Q_1, that is released in the growth medium as the input signal to the Targets. The feedback loop is thus closed through species D which catalyzes the synthesis of the molecule Q_2, whose concentration in the growth medium is interpreted by the Controllers as the system output readout. The proposed topology is such that the input signal (the concentration of Q_1) is an indirect function of the mismatch between the reference signal and the output readout (Q_2 concentration). In what follows, we use the subscript 'e' to denote concentrations of the signaling molecules outside the cells and the subscripts 'c' and 't' to denote their concentrations inside Controller and Target cells respectively.

Note that this specific implementation of the feedback controller architecture can allow re-use of the Controllers to steer different synthetic target populations, as long as the latter are able to establish a bi-directional communication with the former.

Model derivation

The concentration of each species B, C, D, and of the active complex $A{:}Q_2$ in the gene regulatory networks (GRNs) in the Controller and in the Target cells is modeled by Ordinary Differential Equations (ODEs), while spatial diffusion of the signals between the populations is modeled by means of Partial Differential Equations (PDEs).

Activation, or repression, of each species x by its regulator s is governed by a Hill transcription function with dissociation constant K_s and exponent n_s. Species A is simultaneously activated by the signaling molecule Q_2 (forming the active complex $A{:}Q_2$) and repressed by the incoming reference signal $r(t)$ (i.e. is regulated by two independent and competing species), thus the dynamics of complex $A{:}Q_2$ are concisely modeled as the product of the two Hill functions (*25*). Degradation of species x is governed by first order kinetics with corresponding degradation rate γ_x. Moreover, for each species x, characteristic parameters of transcription and translation are embedded in the basal ($\chi_{x,0}$) and maximal (χ_x) activity constants. The model consists of three key ingredients: *i)* a set of equations describing species concentrations inside the Controller cells, *ii)* a set of equations describing species concentrations in the Targets and *iii)* a set of equations to model communications between the two populations. Each of these three components of the consortium model is described below.

Controller population The regulated product B is only produced in response to the active complex $A{:}Q_2$, and is not induced by the individual inactive constituent components of this complex (Figure 1b). The dynamics of the molecular species' concentrations in the Controller population can be written as:

$$\begin{aligned}\frac{d\,[A{:}Q_2]}{dt} = &\left(\chi_{A:Q,r,0} + \chi_{A:Q,r}\frac{{K_r}^{n_r}}{{K_r}^{n_r} + [\mathrm{r}]^{n_r}}\right)\cdot \\ \cdot &\left(\chi_{A:Q,a,0} + \chi_{A:Q,a}\frac{[Q_{2,c}]^{n_q}}{{K_q}^{n_q} + [Q_{2,c}]^{n_q}}\right) - \gamma_{A:Q}[A{:}Q_2],\end{aligned} \tag{1}$$

$$\frac{d\,[B]}{dt} = \chi_{b,0} + \chi_b \frac{[A{:}Q_2]^{n_b}}{{K_b}^{n_b} + [A{:}Q_2]^{n_b}} - \gamma_B[B]. \tag{2}$$

Target population The Target population receives the control input from the Controller cells, by sensing the concentration of molecule Q_1; this catalyzes the synthesis of the species C, which in turn inhibits D (Figure 1b). The dynamics of the molecular species' concentrations in the Targets can be described by:

$$\frac{d\,[C]}{dt} = \chi_{c,0} + \chi_c \frac{[Q_{1,t}]^{n_c}}{{K_c}^{n_c} + [Q_{1,t}]^{n_c}} - \gamma_C[C], \tag{3}$$

$$\frac{d\,[D]}{dt} = \chi_{d,0} + \chi_d \frac{{K_d}^{n_d}}{{K_d}^{n_d} + [C]^{n_d}} - \gamma_D[D]. \tag{4}$$

Communication between the two populations The two populations communicate via two pathways, the first directed from the Controllers towards the Targets (pathway 1, molecule Q_1), whereas the second goes from the Targets to the Controllers (pathway 2, molecule Q_2). The two pathways are symmetrical, indeed for each of them a *sender* and a *receiver* cell population can be defined. Furthermore, the following assumptions are considered modeling the dynamics of Q_1 and Q_2 concentrations:

1. Q_1 and Q_2 are orthogonal so that no cross-talk occurs between different signals;
2. the two molecules diffuse across the cell membranes with the same diffusion coefficient η and are degraded in the cells at the same rate γ_i;
3. the intra-cellular concentration in the sender depends on the rate of production of the molecule (K_Q), on the exchange rate with the extra-cellular environment, and on the degradation inside the cell;
4. the intra-cellular concentration in the receiver is a function of the exchange with the extra-cellular environment and of the degradation inside the cell;

5. the extra-cellular concentrations are functions of the diffusion coefficient Θ in the growth medium, of the exchange between the cells and the extra-cellular environment and of the external degradation rate γ_e;

6. the production of both Q_1 and Q_2 occurs through a synthesis process catalyzed by species B and D respectively, where the substrates required to synthesize both signaling molecules are in excess (*26*). Consequently the synthesis does not have any direct effect (which would only manifest as additional linear dilution terms in Equations (2) and (4)) on the concentrations of the species B and D.

The dynamics of intra-cellular and external concentrations of Q_1 (pathway 1, the Controller is the *sender* and the Target is the *receiver*) can therefore be described as:

$$\frac{d[Q_{1,c}]}{dt} = K_{Q_1}[B] + \eta\big([Q_{1,e}] - [Q_{1,c}]\big) - \gamma_i[Q_{1,c}], \tag{5}$$

$$\frac{d[Q_{1,t}]}{dt} = \eta\big([Q_{1,e}] - [Q_{1,t}]\big) - \gamma_i[Q_{1,t}], \tag{6}$$

$$\frac{\partial[Q_{1,e}]}{\partial t} = \eta\big([Q_{1,c}] - [Q_{1,e}]\big) + \eta\big([Q_{1,t}] - [Q_{1,e}]\big) - \gamma_e[Q_{1,e}] + \Theta\nabla^2[Q_{1,e}]; \tag{7}$$

while those for Q_2 (pathway 2, the Target is the *sender* and the Controller is the *receiver*) as:

$$\frac{d[Q_{2,c}]}{dt} = \eta\big([Q_{2,e}] - [Q_{2,c}]\big) - \gamma_i[Q_{2,c}], \tag{8}$$

$$\frac{d[Q_{2,t}]}{dt} = K_{Q_2}[D] + \eta\big([Q_{2,e}] - [Q_{2,t}]\big) - \gamma_i[Q_{2,t}], \tag{9}$$

$$\frac{\partial[Q_{2,e}]}{\partial t} = \eta\big([Q_{2,c}] - [Q_{2,e}]\big) + \eta\big([Q_{2,t}] - [Q_{2,e}]\big) - \gamma_e[Q_{2,e}] + \Theta\nabla^2[Q_{2,e}]. \tag{10}$$

Note that equations (7) and (10) are PDEs describing the spatio-temporal dynamics of the

concentrations of Q_1 and Q_2 in the extra-cellular environment.

Multicellular feedback control objectives and simulation scenarios

In agreement with the existing literature (*3*, *5*, *7*, *24*), as test-bed goals to validate the effectiveness of the multicellular control approach we consider two different control objectives: *i) set-point regulation*, where the concentration of the output species D has to reach and maintain a specific fixed value; *ii) signal-tracking*, in which the system output D has to track a desired time varying concentration profile. Specifically, we choose the following representative reference signals:

1. *set-point:* $r(t) = 3 \cdot U(t-100)$ (Figure S1;
2. *multi set-point:* $r(t) = U(t-300) + 2 \cdot U(t-700) - U(t-1100)$ (Figure S1b);
3. *trapezoidal:* $r(t) = \alpha \cdot (t-100) \cdot U(t-100) - \alpha \cdot (t-500) \cdot U(t-500) - \alpha \cdot (t-900) \cdot U(t-900) + \alpha \cdot (t-1300) \cdot U(t-1300)$ (Figure S1c);
4. *sinusoidal:* $r(t) = U(t-300) + \sin\left(\frac{2 \cdot \pi \cdot (t-300)}{400}\right)$ (Figure S1d);

where $U(t)$ is the Heaviside step function (*24*) and, α is the slope of the linear increasing and decreasing sections of the trapezoidal signal.

Note that the model predictions show that the steady-state value of the output concentration of D varies non linearly as a function of the reference signal. Specifically, when the reference signal is varied slowly enough to allow the system to settle onto a steady-state for each of the reference levels set between $0\,\mu$M and $3\,\mu$M, the resulting response is significantly nonlinear for concentrations of the control reference signal in the interval $[0, 1.5]\,\mu$M (Figure S2). As done in many nonlinear control schemes (see (*27*)), this information is exploited to perform a nonlinear dynamic inversion, i.e. to calculate the amplitude of the reference signals so that, in the control experiments, the corresponding desired system behavior can be set as the control target (further details available in Supporting Information text).

Next, *in-silico* experiments are carried out by contrasting two alternative numerical implementations of the model. Specifically, the following two scenarios are considered: *i)* an aggregate populations scenario where each of the two populations (Controllers and Targets) are modeled as a single average cell, and *ii)* an agent-based scenario in which individual cells' dynamics are simulated via an agent-based interaction model. While the former implementation assesses the feasibility of the proposed feedback control scheme, along with its robustness and modularity, the latter is used to evaluate the effects of growth, motility, cell density, as well as to assess influence of varying the ratio between Controllers' and Targets' populations and of cell-to-cell variability. Agent-based simulations provide highly accurate simulations of population-level dynamics, while allowing a more realistic representation of the geometry of the host environment where cells grow.

Details of how the model is parameterized are provided in the Methods section and Supporting Information text.

Aggregate populations scenario

The aim of the analysis in this configuration is to test the overall system behavior for different control references (set-point and signal tracking regulation) in the presence of spatial diffusion of the inducer molecules between the Controller and Target cells, quantify the effect of varying the distance between the two interacting cells and assess the modularity of the proposed control implementation.

In this scenario an average Controller cell is assumed to interact with an average Target cell in a mono-dimensional spatial domain at a distance of 20 μm from each other (which is realistic in standard microfluidic devices for bacteria (*18*, *28*, *29*)).

Control results Despite strong nonlinearities in the model and the presence of spatial diffusion of the signaling molecules, the multicellular feedback control strategy is effective in achieving set-point regulation with zero steady-state error, an output overshoot less than

10% of the desired value and a settling time of 150 minutes (Figures 2a, S3a). Similar results, in terms of overshoot and settling time, are exhibited by the multicellular controller when the goal is switched to that of tracking a multi set-point reference signal (Figures 2b, S3b). Tracking of the trapezoidal and the sinusoidal control reference signals is acceptable with a delay in the response of less than 50 minutes and an expected distortion of the reference wave-forms that can be observed in the output (Figures 2c,d and S3c,d).

If communication from the Target to the Controller cell (pathway 2) is absent or suddenly interrupted, the control is said to be in 'open loop' (*24*). In this situation, the Controller population does not receive information on Target's behavior and cannot compensate anymore the deviation between the actual system output (concentration of the species D) and the specific control reference as when this information is available (closed loop). This is confirmed when open loop control is simulated to achieve set-point regulation of a perturbed Target model (see Supplementary Information text for simulation details); the resulting system response presents larger average output steady-state error (4.54%) and, more notably, a considerably larger steady-state standard deviation (42%) computed across the repetitions (Figure S4 a,c) than those achieved when closed loop set-point regulation is simulated on the same perturbed system (average output steady-state error of 0.73% and standard deviation of 33%, Figure S4 b,c). The beneficial effects, on the system's response, of closed loop control are much more obvious in the case of signal tracking regulation. Indeed, when open and closed loop strategies are employed to steer the same perturbed system to track a sinusoidal reference, an even larger variability of the controlled variable is measured in open loop ($CV_{av} = 50.77\%$) than that found in closed loop ($CV_{av} = 33.07\%$) (see Figure S5 and Supporting movie S10).

Robustness analysis In order to assess the robustness of the proposed multicellular control scheme to environmental perturbations, the effect of varying cells' distance in a monodimensional domain is investigated in the case of set-point control. As the distance between

the Controller and the Target cell is linearly increased from 2 μm to 1000 μm, the control action is still effective in regulating the system output to settle onto a fixed concentration, although the concentrations of the input as well as the output signal perceived by the concerned cells are attenuated (see Figure 3). In particular, an increase of cells' distance leads to a significant reduction of the dynamical range of the controlled output when the cells are distant more than 500 μm from each other, from 0.7 μM at a distance of 20 μm, to 0.2 μM at 1000 μm which corresponds to a relative reduction of 71% (Figure 3 b). Moreover, as an effect of the input attenuation for increasing cells' distances, the output overshoot drops until it finally disappears for distances larger than 500 μm (Figure 3 b). The range of distances considered in this scenario is by far the worst case possible in a realistic microfluidic chamber in which cells usually grow much closer; thus, as confirmed by the results achieved via agent-based simulations presented below, the outcome of real microfluidics experiments is expected to exhibit a decay of the communication signals that is negligible with respect to that found in these *in silico* predictions.

Modularity In order to better elucidate how the different modules (populations), within the consortium, contribute to the overall system behavior, the effect of parameter perturbations on the control effectiveness is evaluated by taking set-point regulation as the control goal. Specifically, modularity is assessed in three different instances: *i)* all the parameters are varied, *ii)* only Target's parameters are perturbed in Equations (3), (4), (6) and (9) and *iii)* only Controller's parameters are varied in Equations (1), (2), (5) and (8). In each perturbation case, parameter values are drawn from normal distributions centered in the corresponding nominal values with a standard deviation of 20% (additional details are available in the Methods section). Despite parameter variations, the control action is still effective and robust in driving the average system response (computed across all the repetitions) to settle onto a constant value as required by the specific regulation task (Figure 4a-c). The variability across all the simulation repetitions is higher ($CV_{av} = 41.22\%$) when all parame-

ters are perturbed than in the cases in which only Target's ($CV_{av} = 33.13\%$) or Controller's ($CV_{av} = 21.46\%$) parameters are varied (Figure 4a-c). Notably the Controller behaves well even in the presence of consistent variations of the Target's parameters. This finding is also confirmed by more realistic agent-based *in-silico* experiments as reported below.

Agent-based scenario

The control performance is tested via an agent-based model using the BSim framework (*30*), to account both for cell density and the extracellular environment (i.e., microfluidic devices, which are optimal devices for real-time analysis of engineered bacterial population dynamics (*31*)). Cells represented using BSim's *E. coli* bacterium model, with biologically realistic cell motility (*32*), are inserted in an extra-cellular environment modeled as a rectangular microfluidic chamber with dimensions of $25 \times 20 \times 1$ μm^3, able to host a maximum (subsequently referred to as a density of 100%) of 240 rod-shaped, or 480 spherical cells (*29*).

The chamber is open to the external flow only on both long sides, allowing for diffusion of a reference signal into the chamber and diffusion of the signaling molecules out of the chamber on those sides only (*29*, *33*). Additional model details are provided in Methods section.

Control results The performance of the control feedback strategy in the agent-based scenario is initially tested with the set-point regulation and multi set-point reference signals (Figure S6) as in the case of the aggregate populations simulations. Direct comparison of control outcomes in the agent-based scenario with those measured in the aggregate scenario indicates that settling time is shorter (less than 50 minutes), though at the cost of greater overshoot (up to 18% of the desired set point in the most extreme case) prior to settling.

The agent-based model is further simulated using the other two time-varying reference signals, trapezoidal and sinusoidal, as defined in the previous section (Figure 5a,b). *In-silico* experiments in this scenario confirm the effectiveness of the control strategy as depicted in

Figure 5. The measured output averaged across the Target population (solid lines) tracks the reference signal (dashed lines) closely, for both input wave forms, with a negligible delay. Moreover, the standard deviation of the response across the population is less than 1% of the mean. Qualitatively, the signal-tracking performance is better than in the aggregate model (compare Figures 2 and 5). This is due to the presence in these simulations of a dense population of cells at a comparatively shorter cell-to-cell distance ($< 5\mu$m) than that considered in the aggregate scenario. Animations of the full simulations, along with the quantification of Target population output, are provided in Supporting Movies S11 and S12.

Robustness analysis The multicellular controller is robust to changes in population density in the agent-based implementation (Figure 6). Indeed, the measured output from the Target population maintains a desirable dynamical range even as the total density of cells in the microfluidic chemostat is reduced to 10%. However around this density, and below it, there is a dramatic reduction in the quality of the whole output response, with a clear decrease in the output dynamical range from approximately 0.7 μM to less than 0.1 μM as observed in Figure 6b.

This dependence on density is essentially analogous to what is observed from the results in the aggregate populations scenario shown in Figure 3; an increase in density is associated with a decrease in average cell-to-cell distance. At the maximal cell density, average cell-to-cell distance is approximately 1 cell length (1μm for spherical cells, 2μm for rod-shaped cells), whereas at the minimal cell density (2.5%) of those shown in Figure 6 mean cell-to-cell distance is approximately 10μm. Cell positions were chosen such that a uniform spatial distribution and, therefore, even density of cells was ensured over the course of each experiment.

In a similar manner, the control goal continues to be achieved robustly as the ratio between Controller and Target cells is varied within a fixed global cell density (Figure S7). Even when Controllers' proportion is as low as 10% of the total number of cells, the re-

sponse dynamical range and settling time for the set-point regulation task remain adequate (Figure S7b).

Modularity To assess the effects of changes to the different GRN modules (Controller, Target) in the agent-based scenario, average population response is measured when introducing local (per-cell) perturbations of model parameters in the case of set-point regulation. Parameters within each sub-population of agents (all bacteria, only the Controllers and only the Targets) are perturbed, according to the three instances introduced earlier in the aggregate populations scenario. Moreover, in agent-based simulations cell division and growth are explicitly modeled. As shown in Figure 4d-f the average output behaves adequately. However, the averaged coefficient of variation is 2 percentage points higher than that observed for the aggregate model case when all parameters are varied, and 7 percentage points higher when only Target parameters are perturbed (Figure 4d,e). When only Controllers' parameters are varied, agent-based simulations show relatively little noise across the whole population (Figure 4f). The lower output variability observed in this case arises from the fact that all Targets' parameters are the same (no variation across the Target population), thus only the mean of their response is affected, and an overall lower noise is generated (noise statistics calculated across cells). One representative simulation of those depicted in Figure 4d is provided in Supporting Movie S13.

The results achieved in this scenario confirm what is observed in the aggregate population simulations: the Controller module of the multicellular feedback control strategy remains effective even when attempting to control Targets with a variety of response characteristics, and that consequently a single Controller population could be used to effectively regulate more than one type of Target.

Further *in-silico* experiments where perturbations of the parameters are introduced in the absence of cell growth and division are reported in Figure S8, for the sake of comparison. As shown in Figure S8a-c, better matching between the aggregate cell model predictions and

the agent-based simulation can be obtained by introducing global parameter perturbations across the whole consortium by selecting the same set of perturbed values of the parameters for all cells being simulated. In the absence of perturbations, but presence of growth, a small amount of intrinsic noise resulting from cell division perturbs the system. A full simulation of the consortium's response in the set-point regulation scenario, when growth and division are modeled but perturbations are excluded, is provided in Supporting Movie S14. It is clear that noise due to cell division produces only a small deviation in observed Target population output compared to that induced by direct perturbation of parameters (contrast the histograms in Supporting Movies S13 and S14).

Conclusions

We have described the design and the *in-silico* implementation of a multicellular feedback control strategy within a synthetic cellular consortium of two populations where one acts as a controller trying to regulate the concentration of one molecular species in the other. The two cell populations can communicate via the release of signaling molecules in the shared growth medium and a control reference signal can be provided to the controller population by means of an external inducer molecule.

We have proposed a mathematical representation of the two interacting cell populations considering biologically realistic parameters (*26*, *28*, *33–37*).

The *in-silico* experiments, carried out via two different computational approaches, have shown the effectiveness of the implemented feedback control strategy, despite the parts composing the whole circuit being spatially separated and thus the signals (control input and system output) attenuated by diffusion and propagation in the extra-cellular environment.

The perturbation analysis performed has shown that the proposed multicellular control strategy is robust to large parameter variations. This highlights the inherent modularity of the approach; indeed, according to our *in-silico* results, the same Controller population could be used to efficiently regulate different synthetic Target cells.

We have analyzed the effects of varying cells' spatial distribution (i.e., the distance in the aggregate model scenario and the population density in the agent-based simulations), as well as the relative number of Controller cells while maintaining a constant total consortium size, demonstrating the dependence of the output dynamic range on the consortium density and composition.

Concerning a possible biological implementation of the discussed control strategy, *ad hoc* conceived microfluidic devices can be employed to grow the cells while precisely delivering, from the external environment, only desired perturbations to them (i.e., the control reference r) (*29*, *31*), as well as to regulate/set the ratio of the interacting populations and ensure a proper diffusion of the signaling molecules enabling communication within the consortium. Indeed, it has been demonstrated in (*28*) that microchemostat features (i.e., device topology and structures' aspect ratio) as well as experimental conditions (i.e., pressures and flow rates), can significantly affect the outcome of synthetic circuits whose internal molecular species are activated and/or inhibited as a response to the concentration of signaling molecules diffusing in the growth medium. Moreover, orthogonal or synthetic quorum sensing systems can be exploited to implement communication across the two populations within the consortium (*26*, *37*, *38*).

On the basis of our predictions and of the engineering of microbial communities completed so far (*18*, *19*, *38*, *39*), we believe that the implementation of such a feedback control strategy in living cells could provide a useful tool for the realization of embedded cellular controllers. Synthetic microbial consortia can indeed be employed when a huge amount of biological circuitry is needed to achieve a desired function, with the advantages of reducing the metabolic load in a single cell population, guaranteeing the reliability of the interconnected parts and modules and allowing re-use of engineered populations for different control applications.

Methods

Model parameterization

Model parameterization was carried out according to characteristic ranges of values available in the literature to describe the dynamics of the interacting molecular species of the proposed GRNs (*28*, *33–37*, *40*, *41*). The specific values selected are indicated in Table S1 in the Supporting Information text.

Aggregate populations scenario

Simulations The mathematical model was composed of 8 ODEs and 2 PDEs (Equations (1)–(10)). Simulations were carried out via the MATLAB software package (The MathWorks, Inc.). The PDEs (7) and (10) were discretized over the N points in which the spatial domain was divided by means of the method of finite differences, using a central step discretization. Dirichlet's boundary conditions were imposed in order to solve the resulting system of $2N$ ODEs. The equations resulting from the discretization of (7) and (10) in the space domain and the other ODEs describing the full system were integrated with the ode15s MATLAB solver with a model sampling time of 5 minutes. Initial conditions were calculated by running a simulation (from random initial conditions) with $r(t) = 0$, thus obtaining, as model initial conditions, the steady state values consequently reached by the state variables in the model.

Modularity analysis Parameters' perturbation was carried out by performing 1000 Monte Carlo simulations and picking parameters' values from normal distributions centered on the corresponding nominal value with a standard deviation of 20%. For all the three instances explored (parameter perturbations of either all the cells, or just the Targets or the Controllers) the output average $\mu(t)$ and standard deviation $\sigma(t)$ were calculated across all the simulations performed. These values were then used to calculate the time-averaged coefficient of variation CV_{av} as:

$$CV_{av} = \frac{1}{T}\int_0^T \frac{\sigma(\tau)}{\mu(\tau)} d(\tau), \tag{11}$$

where T is the simulation length in minutes.

Agent-based scenario

Agent-based simulations were performed using the BSim framework (*30*). The full source code of BSim, including the code required to run the agent-based model presented here, is available to download at `https://github.com/BiocomputeLab/bsim`. Further information regarding details of the structure of a BSim simulation, its key features, and implementation details are available at `http://bsim-bccs.sourceforge.net/downloads.html`. System dynamics were simulated both with and without cell growth and division. Simulations of signal tracking (Figures 5 and S6; Supporting Movies S11 and S12), some simulations of parameter perturbations (Figure S8), and all simulations investigating cell density and ratio effects (Figures 6 and S7), were undertaken without considering growth in order to remove any undesired variation in the number of cells present in the simulation, thus reducing unwanted disturbances to the system while the effects of other perturbations were investigated. Simulations in Figure 4d-f and Supporting Movies S13 and S14 included the full model of cell growth and division.

Simulations without cell growth For the experiments without growth, the baseline conditions were as follows. Cells were simulated as individual agents with spherical geometry with diameter of 1 μm, since in the absence of growth and division a full simulation of cell-cell interactions with realistic capsule geometry would have introduced unnecessary extra complexity to the model. Removing cell growth allowed for precise and consistent imposition of cells' density and ratio in the population and therefore isolated the investigation from any consequent effect on the system's response. Initial cell positions were chosen randomly, based on a uniform spatial distribution.

A cuboid geometry was chosen for the spatial simulation domain (referred to also as the chamber, or microfluidic chamber), with dimensions of $25 \times 20 \times 1$ μm^3 in the x, y, and z axes respectively (insets in Figure 5, and Supporting Movies S8 and S9). The chamber was closed on all sides except the two (y-aligned) long sides. Its design (specifically the choice of aspect ratio and boundaries) conformed to those commonly used in microfluidic experiments (*29*), however the simulated chamber was reduced in size compared to those used in *in-vivo* experiments. A chamber of this design is able to host a maximum of approximately 240 rod-shaped, or 480 spherical, cells when cells are present at 100% density. In practice the number of rod-shaped cells may be lower at steady-state due to growth; we discuss this in the following section. The actual number of cells present in each simulation is indicated in the relevant Figures. All boundaries were closed to physical cell passage, so that cells were unable to leave the simulation thus maintaining a constant set ratio of Controllers to Targets and a constant average cell density over time.

Simulations with cell growth In the experiments in which cell growth was taken into account, it was modeled using a per-cell ODE model of rod elongation over time as in (*42*, *43*), with division occurring once the mother cell passed a set constant threshold length. Cell radius was set to a constant value of 1 μm, and initial cell length was set to 2.5 μm with a maximal division threshold of 5 μm. When rod-shaped geometry and growth mechanics are included in the model, cells have on average a greater length than the minimum necessary for 100% density (all rod-shaped cells are guaranteed to always exceed their minimum length). Therefore in practice a lower cell count than that discussed above is observed (approximately 115 cells in total at steady state; Figure S9a). The actual number of cells present in each simulation is indicated in the relevant Figures.

The elongation rate was chosen such that cell division occurred on average after 25 minutes of growth. Upon division, daughter cells' positions and lengths were perturbed by a small amount (randomly chosen, with an amplitude between $\pm 5\%$ of their length) at the

location of division in order to break axial symmetry (*42*). Cells' positions were seeded at $t = 0$ such that all cells were present at a uniform density, with Controllers occupying half of the chamber and Targets the other half. This was in order to increase the chances of maintaining the desired 1:1 ratio of Controllers to Targets throughout the simulation, in spite of stochastic growth effects.

The chamber's physical boundaries were modified such that the two long sides, through which signal diffusion occurred, would permit passage of cells in order to model the removal of cells via flow as occurs in real microfluidic experiments (*29*). A lateral force corresponding to a typical flow rate of 250 μms^{-1} was then applied to any cells passing outside these two open boundaries. Visualizations of full simulations with growth and cell division are shown in Supporting Movies S13 and S14.

GRNs and signaling PDEs modeling the communication across the two different cell populations (Equations (7) and (10)) were simulated using a finite volume discretization, with diffusion and degradation rates specified in Table S1. Reflective boundary conditions were applied to the discretized equations for the four closed sides of the domain. Dirichlet boundary conditions were applied at the two open sides, with a constant zero external concentration representing the diffusion-dominated flow out of the microfluidic chamber.

Internal GRNs were implemented in individual cells using the ODEs defined in Equations (1) and (2) for the Controllers, and Equations (3) and (4) for the Target cells. The intracellular embedded GRN equations were coupled to the discretized concentrations of Q_1 and Q_2 at the position of each cell, and were solved using a fixed-step, 4^{th} order, Runge-Kutta method. Initial conditions for all GRN ODEs were set to the same values (uniformly zero) as in the aggregate model scenario; simulations were run for a 100 minutes 'burn-in' period in order for the resulting transient to settle, prior to each *in-silico* experiment. All plotted results assume the convention that time $t = 0$ in a given experiment corresponds to the end of this 'burn-in' period.

Modularity analysis Two types of perturbation were applied in the agent-based scenario for modularity analysis: a *local* and a *global* perturbation. For both cases, the perturbation was applied either to all agents, only to the Target or only to the Controller agents. In the local case, the parameters of each agent were independently drawn from normal distributions centered on the parameter's corresponding nominal value and with a standard deviation equal to 20% of each nominal value. In the global case, the same parameter values drawn from distributions as above were applied to all agents in the population of interest.

As in the case of modularity analysis in the aggregate population scenario, the output average $\mu(t)$ and standard deviation $\sigma(t)$ were calculated across all the simulations performed for each single series of perturbations, and the time-averaged coefficient of variation CV_{av} was calculated according to Equation (11).

Supporting Information

The material supplied as supporting information is organized as follows:

- Supporting Information text
- Supporting Figures S1-S9
- Supporting Movies S10-S14

Author Contribution

GF and AM contributed equally as first authors. LM and MdB contributed equally to this work. GF, AM, FA, CG, NJS, LM and MdB conceived the idea of the study; FA and NJS contributed to the biological implementation of the feedback controller. GF and AM derived the mathematical model of the system and generated and analyzed the data. GF implemented the aggregate population scenario and AM the agent-based study. GF, AM, FA, CG, NJS, LM and MdB wrote the paper.

Funding sources

This work was supported by BrisSynBio, a BBSRC/EPSRC Synthetic Biology Research Centre, grant number: BB/L01386X/1.

References

1. Laplante, M., and Sabatini, D. M. (2012) mTOR signaling in growth control and disease. *Cell 149*, 274–293.

2. Yi, T.-M., Huang, Y., Simon, M. I., and Doyle, J. (2000) Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl. Acad. Sci. U. S. A. 97*, 4649–4653.

3. Milias-Argeitis, A., Summers, S., Stewart-Ornstein, J., Zuleta, I., Pincus, D., El-Samad, H., Khammash, M., and Lygeros, J. (2011) In silico feedback for in vivo regulation of a gene expression circuit. *Nat. Biotechnol. 29*, 1114–1116.

4. Toettcher, J. E., Gong, D., Lim, W. A., and Weiner, O. D. (2011) Light-based feedback for controlling intracellular signaling dynamics. *Nat. Methods 8*, 837–839.

5. Uhlendorf, J., Miermont, A., Delaveau, T., Charvin, G., Fages, F., Bottani, S., Batt, G., and Hersen, P. (2012) Long-term model predictive control of gene expression at the population and single-cell levels. *Proc. Natl. Acad. Sci. U. S. A. 109*, 14271–14276.

6. Menolascina, F., Fiore, G., Orabona, E., De Stefano, L., Ferry, M., Hasty, J., di Bernardo, M., and di Bernardo, D. (2014) In-vivo real-time control of protein expression from endogenous and synthetic gene networks. *PLoS Comput Biol 10*, e1003625.

7. Fiore, G., Perrino, G., di Bernardo, M., and di Bernardo, D. (2015) In Vivo Real-Time Control of Gene Expression: A Comparative Analysis of Feedback Control Strategies in Yeast. *ACS Synth. Biol.*

8. Fracassi, C., Postiglione, L., Fiore, G., and di Bernardo, D. (2015) Automatic Control of Gene Expression in Mammalian Cells. *ACS Synth. Biol.*

9. Dunlop, M. J., Keasling, J. D., and Mukhopadhyay, A. (2010) A model for improving

microbial biofuel production using a synthetic feedback loop. *Syst. Biol. Synth. Biol. 4*, 95–104.

10. Zhang, F., Carothers, J. M., and Keasling, J. D. (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol. 30*, 354–359.

11. Danino, T., Prindle, A., Kwong, G. A., Skalak, M., Li, H., Allen, K., Hasty, J., and Bhatia, S. N. (2015) Programmable probiotics for detection of cancer in urine. *Sci. Transl. Med. 7*, 289ra84–289ra84.

12. Bloom, R. J., Winkler, S. M., and Smolke, C. D. (2015) Synthetic feedback control using an RNAi-based gene-regulatory device. *J. Biol. Eng. 9*, 1.

13. Briat, C., Gupta, A., and Khammash, M. (2016) Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Bimolecular Networks. *Cell Systems 2*, 15–26.

14. Hsiao, V., de los Santos, E. L., Whitaker, W. R., Dueber, J. E., and Murray, R. M. (2014) Design and implementation of a biomolecular concentration tracker. *ACS Synth. Biol. 4*, 150–161.

15. Kwok, R. (2010) Five hard truths for synthetic biology. *Nature 463*, 288.

16. Weiße, A. Y., Oyarzún, D. A., Danos, V., and Swain, P. S. (2015) Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl. Acad. Sci. U. S. A. 112*, E1038–E1047.

17. Macia, J., Posas, F., and Solé, R. V. (2012) Distributed computation: the new wave of synthetic biology devices. *Trends Biotechnol. 30*, 342–349.

18. Chen, Y., Kim, J. K., Hirning, A. J., Josić, K., and Bennett, M. R. (2015) Emergent genetic oscillations in a synthetic microbial consortium. *Science 349*, 986–989.

19. Brenner, K., You, L., and Arnold, F. H. (2008) Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol. 26*, 483–489.

20. Teague, B. P., and Weiss, R. (2015) Synthetic communities, the sum of parts. *Science 349*, 924–925.

21. Del Vecchio, D., Ninfa, A. J., and Sontag, E. D. (2008) Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol. 4*, 161.

22. Jayanthi, S., Nilgiriwala, K. S., and Del Vecchio, D. (2013) Retroactivity controls the temporal dynamics of gene transcription. *ACS Synth. Biol. 2*, 431–441.

23. Del Vecchio, D. (2015) Modularity, context-dependence, and insulation in engineered biological circuits. *Trends Biotechnol. 33*, 111–119.

24. Aström, K. J., and Murray, R. M. *Feedback systems: an introduction for scientists and engineers*; Princeton university press, 2010.

25. Alon, U. *Feedback systems: an introduction for scientists and engineers*.

26. Balagaddé, F. K., Song, H., Ozaki, J., Collins, C. H., Barnet, M., Arnold, F. H., Quake, S. R., and You, L. (2008) A synthetic Escherichia coli predator–prey ecosystem. *Mol. Syst. Biol. 4*, 187.

27. Khalil, H. K. *Nonlinear control*; Prentice Hall, 2014.

28. Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J. (2010) A synchronized quorum of genetic clocks. *Nature 463*, 326–30.

29. Mondragón-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L., and Hasty, J. (2011) Entrainment of a population of synthetic genetic oscillators. *Science 333*, 1315–1319.

30. Gorochowski, T. E., Matyjaszkiewicz, A., Todd, T., Oak, N., Kowalska, K., Reid, S., Tsaneva-Atanasova, K. T., Savery, N. J., Grierson, C. S., and di Bernardo, M. (2012) BSim: an agent-based tool for modeling bacterial populations in systems and synthetic biology. *PloS one 7*, e42790–e42790.

31. Ferry, M., Razinkov, I., and Hasty, J. (2011) Microfluidics for synthetic biology: from design to execution. *Methods Enzymol. 497*, 295.

32. Berg, H. C. (2000) Motile behavior of bacteria. *Phys. Today 53*, 24–30.

33. Mina, P., di Bernardo, M., Savery, N. J., and Tsaneva-Atanasova, K. (2013) Modelling emergence of oscillations in communicating bacteria: a structured approach from one to many cells. *J. R. Soc., Interface 10*, 20120612.

34. Stewart, P. S. (1998) A review of experimental measurements of effective diffusive permeabilities and effective diffusion coefficients in biofilms. *Biotechnol. Bioeng. 59*, 261–272.

35. De Jong, H. (2002) Modeling and simulation of genetic regulatory systems: a literature review. *J. Comput. Biol. 9*, 67–103.

36. Fiore, G., Menolascina, F., di Bernardo, M., and di Bernardo, D. (2013) An experimental approach to identify dynamical models of transcriptional regulation in living cells. *Chaos 23*, 025106.

37. Scott, S. R., and Hasty, J. (2016) Quorum Sensing Communication Modules for Microbial Consortia. *ACS Synth. Biol.*

38. Song, H., Ding, M.-Z., Jia, X.-Q., Ma, Q., and Yuan, Y.-J. (2014) Synthetic microbial consortia: from systematic analysis to construction and applications. *Chem. Soc. Rev. 43*, 6954–6981.

39. Macia, J., Manzoni, R., Conde, N., Urrios, A., de Nadal, E., Solé, R., and Posas, F.

(2016) Implementation of Complex Biological Logic Circuits Using Spatially Distributed Multicellular Consortia. *PLOS Comput Biol 12*, e1004685.

40. Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S., and Hasty, J. (2008) A fast, robust and tunable synthetic gene oscillator. *Nature 456*, 516–9.

41. Lies, M., and Maurizi, M. R. (2008) Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in Escherichia coli. *J. Biol. Chem. 283*, 22918–29.

42. Jönsson, H., and Levchenko, A. (2005) An Explicit Spatial Model of Yeast Microcolony Growth. *Multiscale Model. Simul. 3*, 346–361.

43. Cho, H., Jönsson, H., Campbell, K., Melke, P., Williams, J. W., Jedynak, B., Stevens, A. M., Groisman, A., and Levchenko, A. (2007) Self-organization in high-density bacterial colonies: efficient crowd control. *PLoS Biol. 5*, e302.

Figure Legends

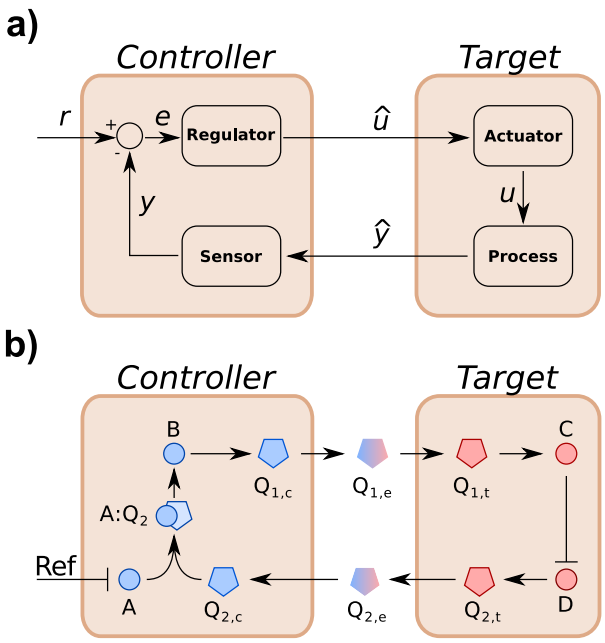

Figure 1: **multicellular feedback control design.** (a) The Controller population contains the sensor to retrieve information on the process status in the Targets, and the regulator to provide inputs to the Target cells according to the control error e computed by the comparator as the difference between the reference r and the measured output y. (b) Abstract biological implementation of the proposed feedback control strategy where molecular species contained in the Controller cell implement the comparing, sensing and regulating functions, whilst species C and D in the Target represent respectively the actuator and the process to be controlled. The roles of the actual control signal and the process output readout are played by the two signaling molecules Q_1 and Q_2. Circles represent internal molecular species and polygons signaling molecules.

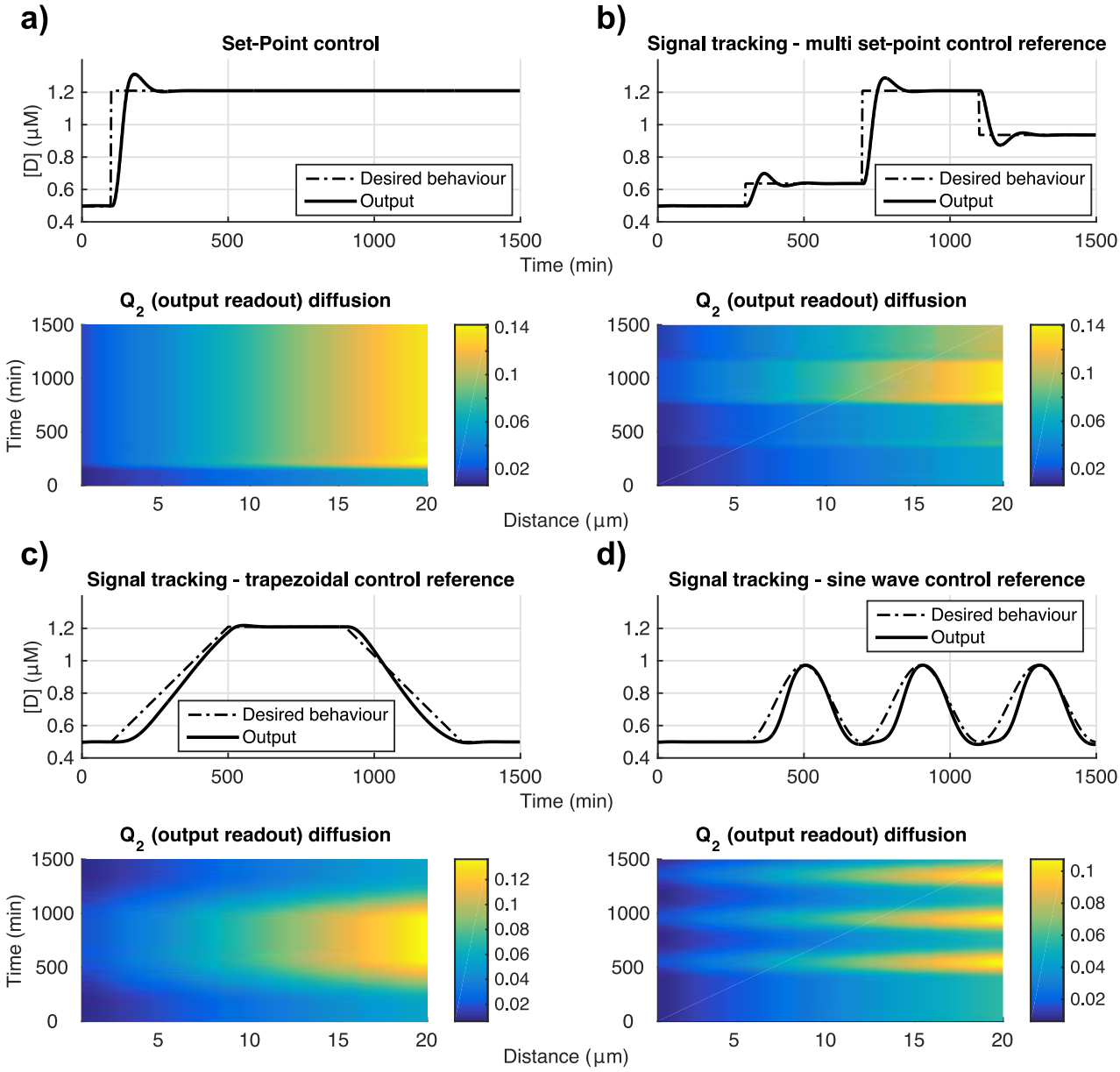

Figure 2: **Aggregate populations scenario: control results for the set-point and the tracking control tasks.** System dynamical response in the case of set-point regulation (a) and tracking of time varying reference signals (b-d); the concentration of the output species D (solid line) and the desired output (dashed line) are plotted for each of the control reference signals. (Insets) Concentration of the signaling molecule Q_2 over the time, indicated in μM, in the extra-cellular environment is depicted for each of the control tasks (Controller cell located at $x_C = 0\mu$m and Target cell located at $x_T = 20\mu$m).

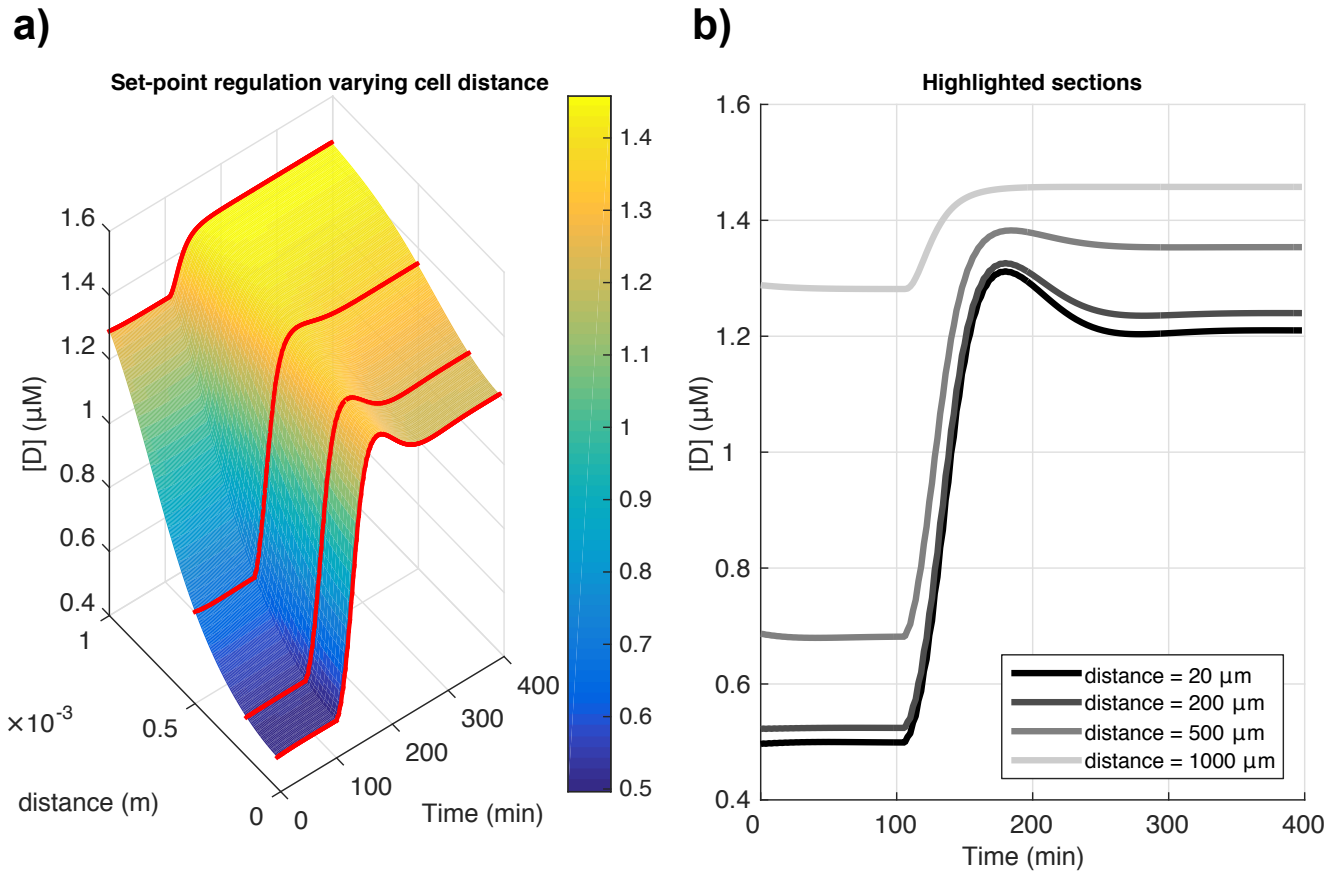

Figure 3: **Aggregate populations scenario, effect of varying cell distance on set-point regulation.** (a) Surface plot of the trajectories obtained for the set-point regulation task when linearly increasing the distance between the Controller and the Target cells from 2 μm to 1000 μm. Red lines highlight the trajectories depicted in panel (b). (b) Species D concentration for distances between Controller and Target cells of 20 μm, 200 μm, 500 μm and 1000 μm.

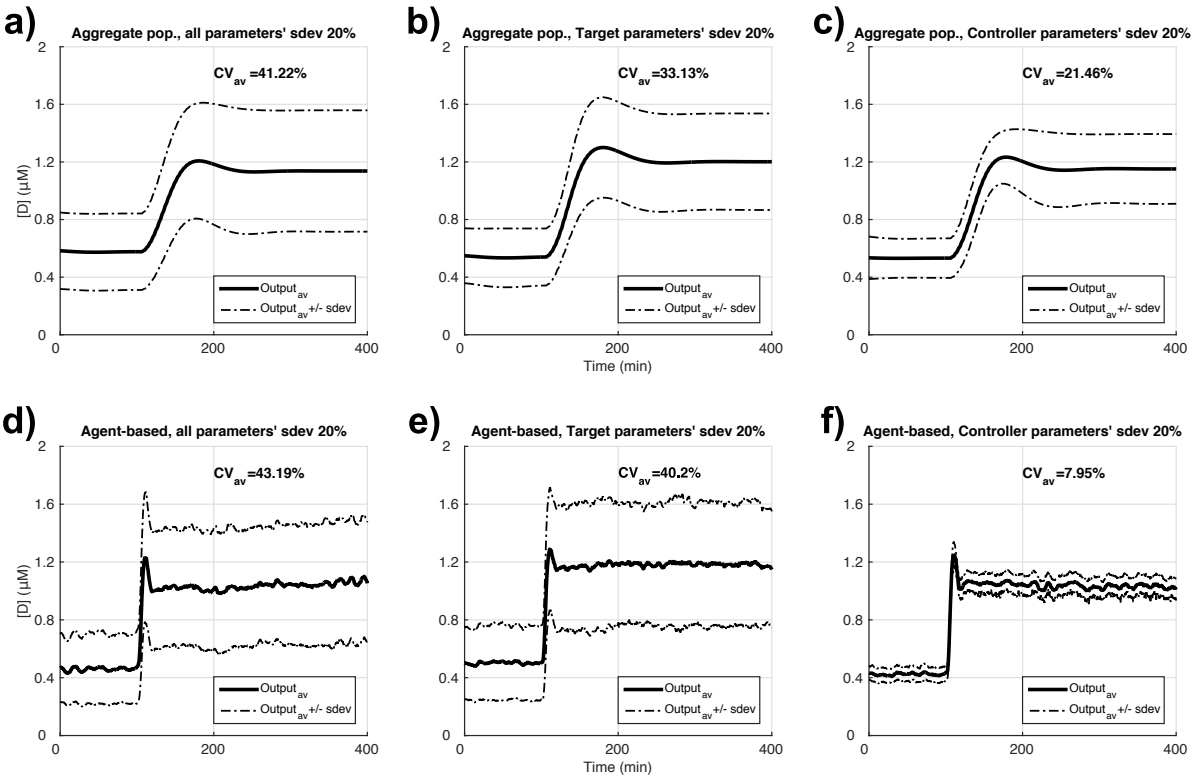

Figure 4: **Modularity analysis, aggregate populations and agent-based scenarios**. (a-c) System dynamical response in the aggregate populations scenario in the case of set-point regulation as a result of a Monte Carlo simulation approach when all the model parameters (a), only Target cell parameters (b) and only Controller cell parameters (c) are perturbed as indicated in the text. (d-f) System dynamical response in the agent-based scenario, when every agent's parameters from complete populations (d), only the Target cells (e) and only Controller cells (f) are sampled individually as indicated in the text. Per-cell growth and division was included in the model, resulting in a stable number of approximately 115 cells in the chamber on average (Figure S9a). Averaged output (solid line) and averaged output $+/-$ the standard deviation (dashed line) across all the simulations performed for each perturbation type are plotted. In the aggregate scenario (a-c), statistics in each panel are computed from 1000 Monte Carlo simulations; in the agent-based scenario (d-f), statistics in each panel are computed from 10 simulations.

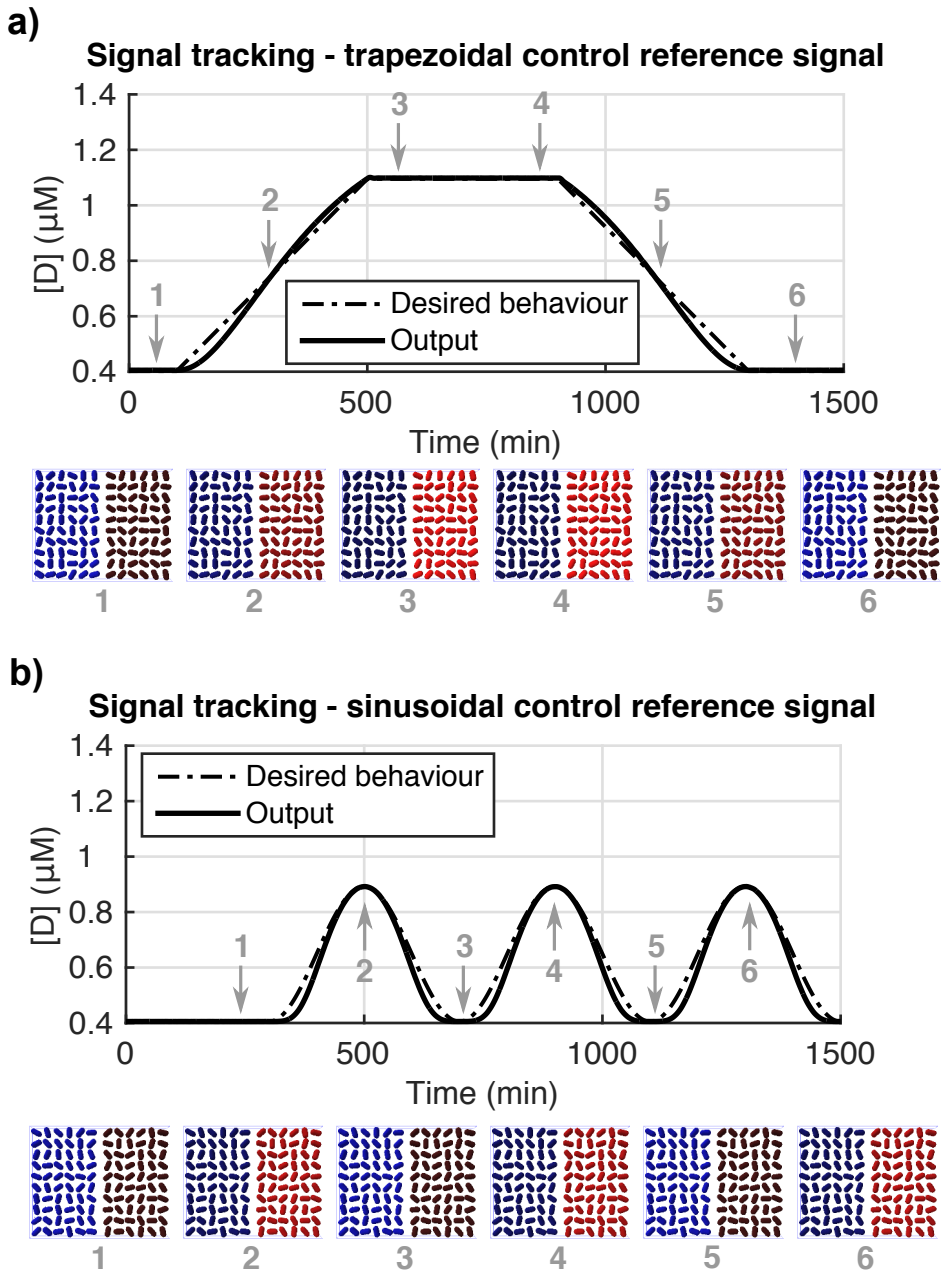

Figure 5: **Agent-based scenario control results.** Agent-based system's response to a trapezoidal reference signal (a) and sinusoidal control reference (b); the average output of the Target population (solid line) and the desired output (dashed line) are plotted for each of the control tasks. (Insets) Snapshots from simulations corresponding to the time points indicated by arrows in panels (a) and (b). Cells are drawn as capsules, with Controllers colored Blue and Targets colored red; the color intensity is directly proportional to $[B]$ in the Controllers and $[D]$ in the Targets. The purple box indicates the extents of the simulation domain. Simulations were performed with 120 rod-shaped cells, without growth and division. The full simulations of the agent-based model, and corresponding quantification, are provided in Supporting Movies S11 and S12.

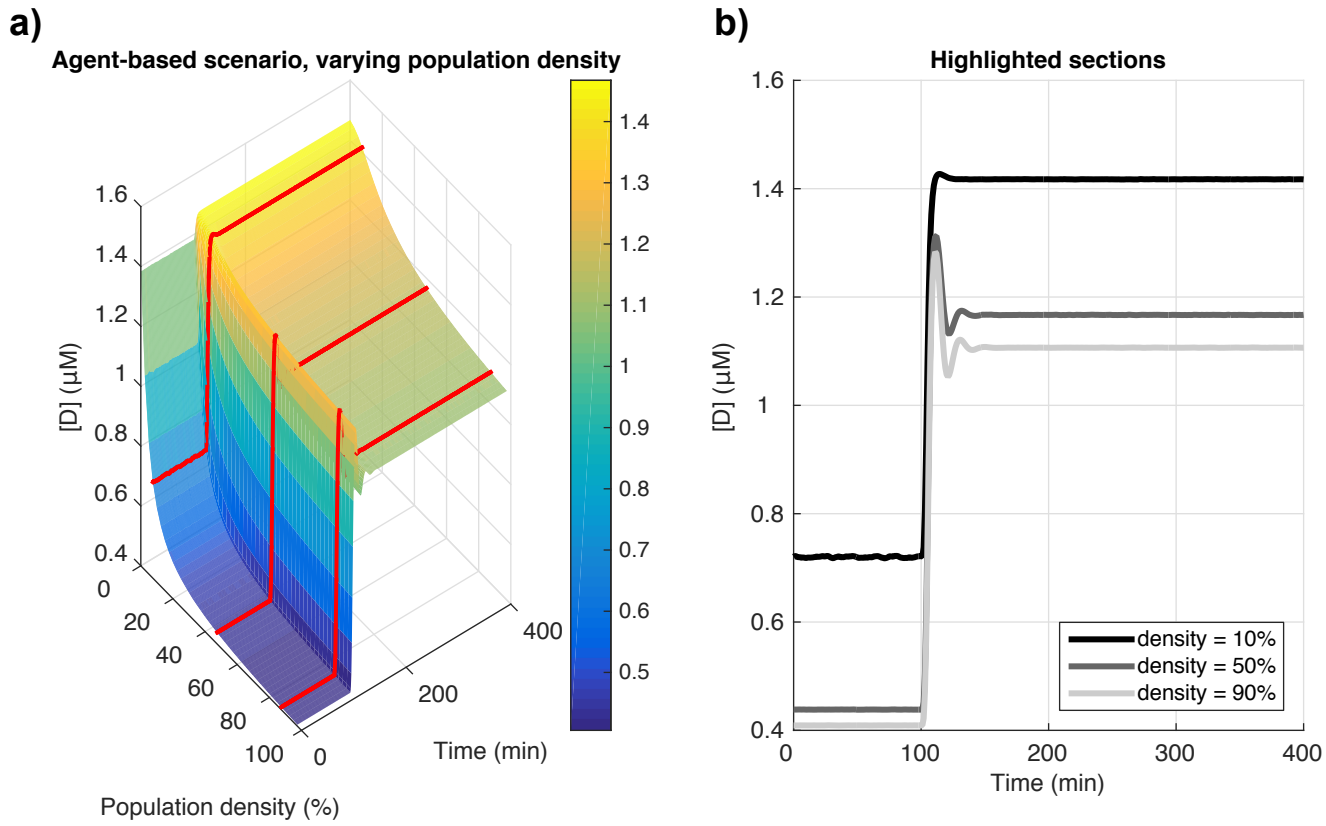

Figure 6: **Agent-based scenario, set-point regulation varying total cell density.** (a) Set-point regulation dynamical response obtained in the agent-based scenario when the number of total cells in the simulation is varied between a density of 2.5% up to a nominal 100% of 480 spherical cells. The ratio of Controllers to Targets was maintained at 1:1. (b) Concentrations of species D for population densities of 10%, 50% and 90%, corresponding to the red lines in panel (a). Simulations were performed without growth and division.